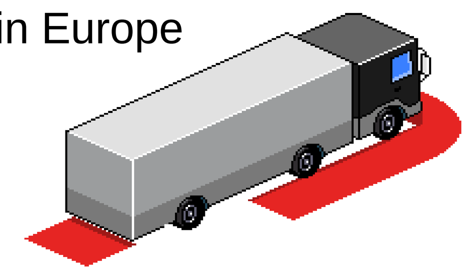
Kristof Van Beeck and Toon Goedemé
EAVISE, Technology Campus De Nayer, KU Leuven, Belgium
{kristof.vanbeeck, toon.goedeme}@kuleuven.be

Embedded & Artificially intelligent VISion Engineering

Efficient Multiclass Object Detection: Detecting Pedestrians and Bicyclists in a Truck's Blind Spot Camera

Problem & Goal

The Blind spot zone:
- Zone around truck in which driver has no/limited view
- Each year approximately 1300 casualties in Europe



Existing systems:
- 2 types: active and passive

- **Passive systems** require truck driver interpretation
  e.g. mirrors, camera system → not adjusted or used correctly

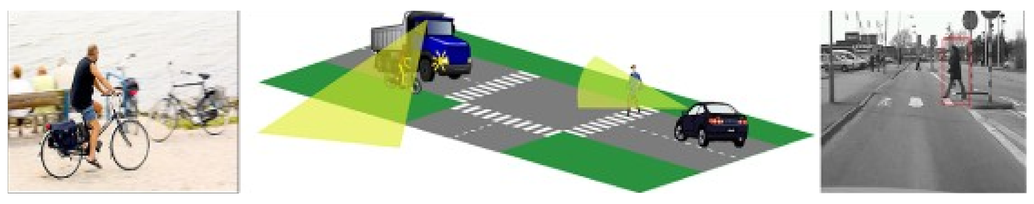
- **Active systems** automatically generate an alarm
  e.g. ultrasonic or infrared systems → cannot distinguish persons vs static objects (traffic signs)

Goal: **Develop a system that automatically detects vulnerable road users (VRUs) in blind spot camera systems**

Advantages of such a system:
- Active driver independent
- Automatic detection of vulnerable road users
- Automatic warning on monitor
- Easy to integrate in existing camera systems
- Reliable, provides implicit interpretation

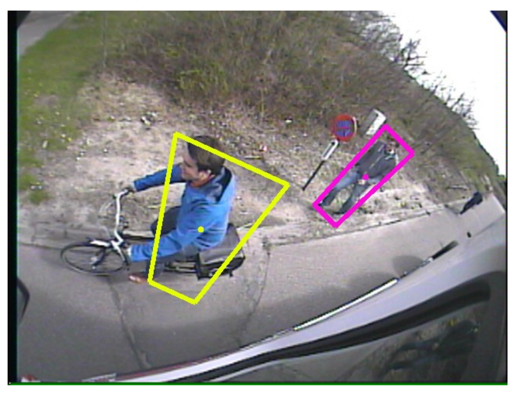
Challenges:
- Hard **real-time**, motion blur, **highly dynamical background**
- Vulnerable road users are a very **diverse class**:

- Blind spot camera: **wide-angle lens & specific viewing direction**

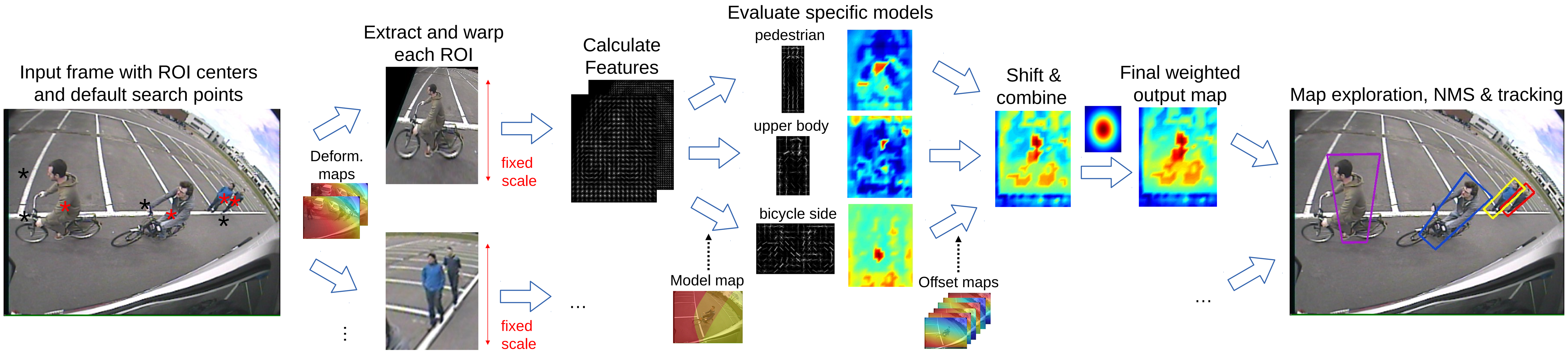


We target **pedestrians** and **bicyclists**
(involved mostly in these type of accidents)

Output detections



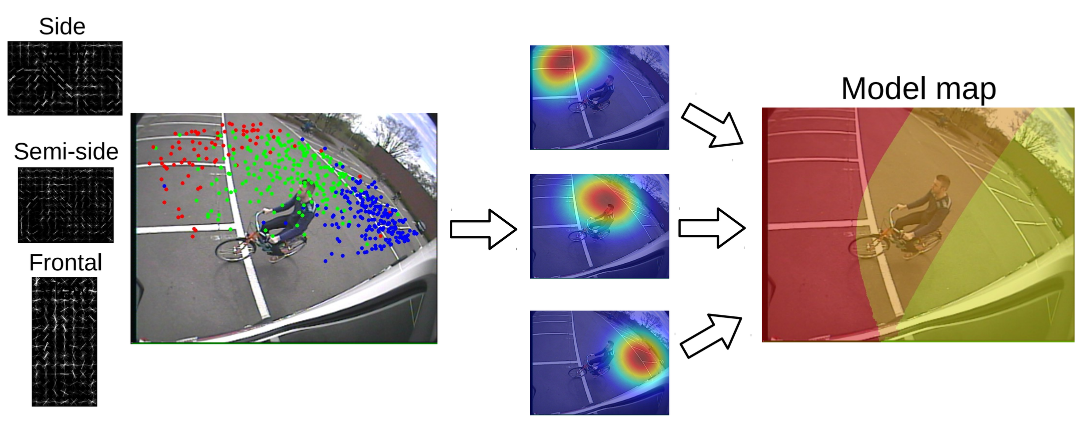
Algorithmic Approach



Algorithmic approach overview:
- Traditional object detectors: *search over full scale-space pyramid*: not feasible here
- VRUs appear under various rotations and scales
- Exact **transformation** only depends on **position in image**
- First **warp each ROI** to an undistorted, upright and fixed height image patch
- Run **multiple detection models** with feature sharing
- Finally, **combine probability maps** and integrate in **tracking-by-detection framework**

1. Extract and warp patches [15]:
- Assume **flat ground plane**, planar objects
- Modeled as **perspective transformation**
- Known after one-time calibration

2. Object detection pipeline:
- Evaluation on a **single scale only**: motivation to use the **cascaded DPM** detector [10] as baseline
- Extract features for each patch, and run **three detection models**: pedestrian, upper body and bicycle component
- Select **best bicycle component** based on position in the image:

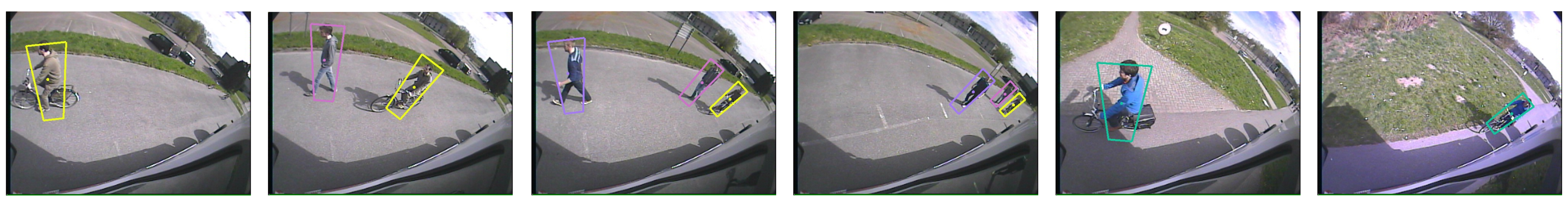


3. Combine probability maps:
- Shift maps with trained offset and combine probability maps using:

$$P_{final}(\mathbf{x}) = \max_{\substack{\mathbf{x} \\ i \in \{1,2,3\}}} \left(P_i(\mathbf{x}) - d_i(\mathbf{x})\right) + G(\mathbf{x}) \qquad G(\mathbf{x}) = \alpha \left[e^{\left(-\frac{\mathbf{x}^2}{2\sigma^2}\right)} - 1 \right]$$

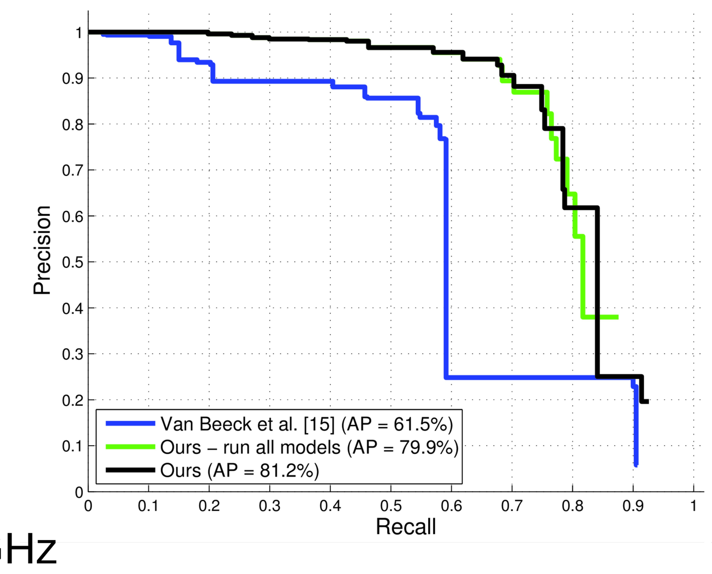
4. Map exploration, NMS and tracking:
- Integrate in **tracking-by-detection framework**:
  - **Default search points** at strategic locations
  - **Threshold probability maps** to extract local maxima and perform NMS
  - Instantiate **Kalman filter** for each new detection
  - Constant **velocity model**, predict **future location**
  - **Verify** for each detection if **track exists** within radius based on scale, otherwise start **new track**

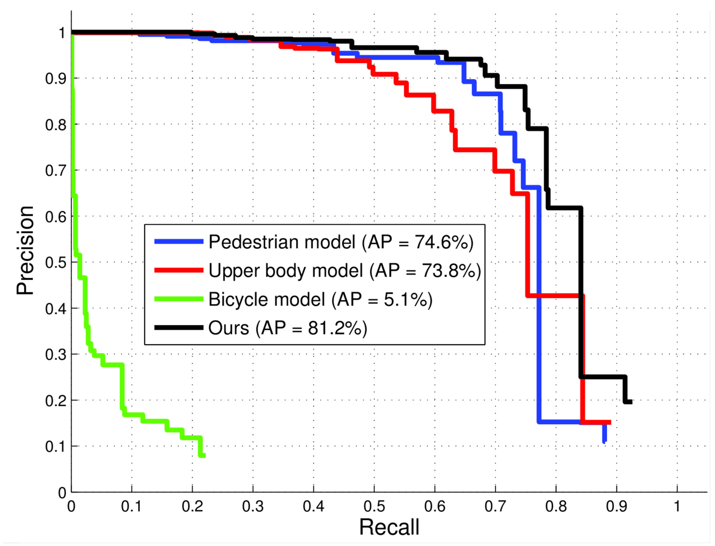
Experiments & Results

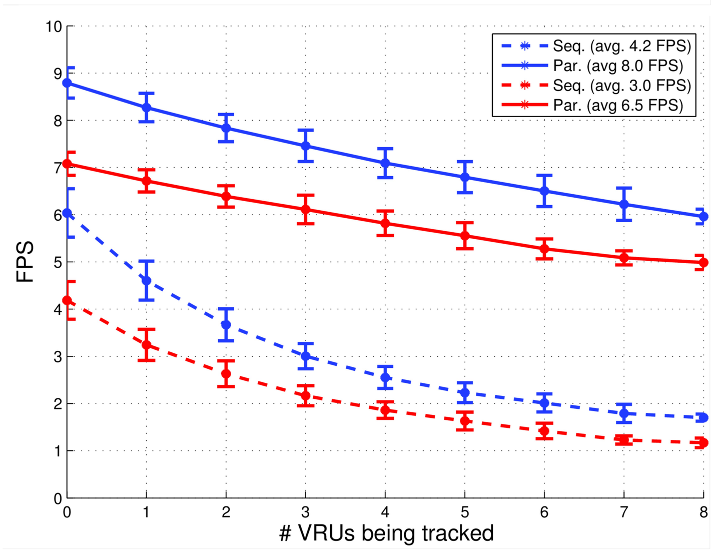


Frame 165 | Frame 215 | Frame 265 | Frame 315 | Frame 490 | Frame 575

We performed extensive experiments:
- Commercial blind spot camera (Orlaco115) on genuine truck (Volvo FM12)
- Simulated **seven** different **dangerous** blind spot situations
- Total test set about **5000 frames**: 640x480 with 15 fps: 3600 pedestrians and 2400 bicyclists labeled
- Both single core and parallel Implementation, Matlab code with C and OpenCV, Intel Xeon E5 CPU at 3.1 GHz
- We achieve **excellent accuracy** (AP = 81.2%) at an average of **8 fps** (Matlab proof-of-concept implementation)







Conclusion

- We presented a multiclass object tracking framework targeting a specific application: detection and tracking of pedestrians and bicyclists in the blind spot camera of a truck
- Using an **integration** of a **warping window approach** with an **efficient detection scheme** where we only run specific viewpoint detectors
- We achieve **excellent accuracy results**, while keeping the **computational complexity** adequate for practical applications

[15] K. Van Beeck and T. Goedemé, Real-Time Pedestrian Detection in a Truck's Blind Spot Camera (ICPRAM 2014) [10] P. Felzenszwalb, R. Girschick and D. McAllester, Cascade Object Detection with Deformable Part Models (CVPR 2010)